Biomarkers: are realism and control mutually exclusive in integrated pollution assessment?

Wilson, J.G.[1], McHugh, B[2]. & Giltrap, M[1].

[1]Zoology Dept., TCD, Dublin 2, Ireland; [2]Marine Institute, Oranmore, Co. Galway, Ireland.
Email: jwilson@tcd.ie

Abstract

The conventional view of pollution monitoring is that any choice is a trade-off between realism and precision, as the control over confounding variables decreases with the increasing degree of organization of the test system. Dublin Bay is subject to considerable anthropogenic pressures and there have been many attempts to quantify the status of the system at organizational levels from DNA strand breaks (Comet) to the system itself (Ecological Network analysis, ENA).

Using Dublin Bay as an example, the data show there was considerable variability at all levels of organization. At intracellular level, Lysosome Membrane Stability (LMS, assessed by Neutral Red Retention, NRR) varied almost 4-fold with season and individual condition, while the community level AZTI Marine biotic Index (AMBI) had a similar range within a single, supposedly homogeneous, site. Overall, there was no evidence that biomarkers of the lower levels of organisation reduced the variability of the measure, despite the extra control over influencing variables, nor was there any evidence that variability was additive at higher levels of organisation.

This poses problems for management, especially given the fixed limits of Ecological Quality Standards (EQSs). Clearly while the integrated approach to pollution monitoring does offer the potential to link effects across the organizational range, it should also be possible to improve their capability by widening the database for reference values, particularly at the higher level of organization, and by process models, including the confounding variables found in the field, for those at lower level.

Introduction.

The recent imperatives imposed by the Water Framework Directive (WFD, 2000/60/EC) and the Marine Strategy Framework Directive (MSFD, 2008/56/EC) with their requirements not just to categorise ecological or environmental status but to achieve at least good status within a defined time-frame have focused attention on the means by which such assessments are made.

The conceptual model of Pearson and Rosenberg (1978) brought together much of the work that had been done to date on pollution changes at community level, and there have been many attempts to develop an index which might reliably summarise the degree of impact. These attempts have included spatial integration formulae (Leppakoski 1977, Jeffrey et al. 1985), mathematical models such as the log-normal distribution (Gray and Mirza 1979), and a variety of diversity/dominance measures such as the Shannon-Weiner distribution which has answered so well in fresh waters before being refined into the current species-based AMBI and Biotic Coefficient (BC) (Borja et al. 2003).

However this traditional approach is costly in both in terms of resources required and in time, not just to carry out the requisite sampling and analysis but also in terms of the lag or inertia in such large systems. An additional complication is imposed by the need to account for natural variability since few if any of the stressors can be controlled (Irvine 2004). Accordingly other measures have been proposed by which the status of a system might be evaluated in a more timely and cost-effective fashion by measuring the performance of a component of the system, rather than the whole thing itself. In addition, these results should be less variable since more of the external variables can be controlled. In addition such components could be selected for their response to specific stressors or contaminants, such as metallothionein (MT) for metals (Viarengi et al. 1997), various elements of the cytochrome P-450 system (Porte et al. 1991) for xenobiotics and imposex for tributyl tin (TBT) (Gibbs et al. 1987). The disadvantage of using components of the system is that the effects at system level may be masked by various homeostatic mechanisms in the system – for instance where one component may be able to compensate for decreased performance in another – or by intrinsic problems such as hormesis in the component response itself (Stebbing, 1979). As a consequence, the choice of monitoring is often depicted as a compromise between realism at community or system level and speed and sensitivity at lower levels of organization as depicted in Figure 1.

A further advantage of measurements at lower levels of organization is that some at least of the confounding variables can be controlled, which means that a more specific response is being measured and unwanted sources of error and variability are eliminated.

The current recommendation from the International Council for the Exploration of the Sea (ICES) Working Group on Biological Effects of Contaminants (Davies and Vethaak 2012) is that a suite of indices be employed from a range of organizational levels to obtain as complete an assessment as possible.

In this study, we present the results from a series of indices of status in Dublin Bay, Ireland, specifically to test whether a) the assessments are consistent among themselves; and b) whether in fact indices at lower levels of organization are less variable.

Materials and Methods.

The test site, Dublin Bay is shown in Figure 2, along with the locations mentioned in the text. Dublin Bay is a shallow, largely sandy system, dominated by various *Venus* (*sensu* Thorson, 1957) communities, and is surrounded on three sides by the conurbation of the city of Dublin. The major riverine input is the River Liffey whose estuary hosts Ireland's largest shipping port and which also receives the city's sewage discharge. Following substantial upgrading in the 1990s, the effluent now receives secondary treatment. Overall, the ecological condition of the Bay has been graded as 'moderate' with serious problems confined to the Tolka estuary and the quayed section of the Liffey (EPA, 2010).

The range of indices is shown in Table 1, along with the level of ecological organization, the type of response measured (as the basis for the index) and the location of the sampled sites in Dublin Bay.

Table 1. Indices tested, level of organization, response measured and location with reference(s) where applicable: see text for references.

Biomarker	Level of Organisation	Response measured	Location
Ecological Network Analysis (ENA)	Ecosystem	Trophic structuring	Intertidal area Whole system
Biological Quality Index (BQI)	Ecosystem	Extent of faunal impoverishment	Intertidal Tolka estuary
Pollution Load Index (PLI)	Ecosystem	Degree of sediment contamination	Intertidal Tolka estuary
AMBI BC Biotic Coefficient	Community	Balance of pollution-sensitive species	Sub-tidal
Shannon-Weiner (H')	Community	Macrofaunal diversity	Sub-tidal
Species Number (S)	Community	Macrofaunal diversity	Sub-tidal
Stress-On-Stress (SOS)	Population	*M. edulis* survival on emersion	Intertidal Tolka estuary
Whole sediment toxicity	Population	LD_{50} and burrowing response of *A. marina, C. volutator*	Intertidal Tolka estuary
Sediment elutriate toxicity	Population	LD_{50} (*T. battaglia, S. costatum, V. fischeri*)	Intertidal Tolka estuary
Sediment porewater toxicity	Population	LD_{50} (*T. battaglia, S. costatum, V. fischeri*)	Intertidal Tolka estuary
Fish liver pathology	Population	Incidence of disease	Sub-tidal
SFG and components (R, CR)	Individual	*M. edulis* individual-level energy budget	Intertidal Tolka estuary
Condition Index	Individual	*M. edulis* flesh weight per unit shell length	Intertidal Tolka estuary

Imposex	Individual	*N. lapillus*: male characteristics in female	Intertidal
Intersex	Individual	*L. littorea*: female abnormality	Intertidal
Ferric Reducing Ability of Plasma (FRAP) assay	Cell	*M. edulis:* cellular antioxidant level	Intertidal Tolka estuary
Ethoxyresorufin-O-deethylase (EROD) assay	Cell	Flatfish: Enzyme induction by chemicals	Sub-tidal (fish)
NRR lysozyme	Cell	*M. edulis*: Lysosome latency (self destruction)	Intertidal Tolka estuary
ALP	Cell	*M. edulis*: Measure of osteogenic differentiation	Intertidal Tolka estuary
Vitellogenin	Cell	*M. edulis*: Egg-yolk precursor in males	Intertidal Tolka estuary
AChE	Cell	*M. edulis* Enzyme activity (muscle, brain)	Intertidal Tolka estuary
MT	Cell	*M. edulis* Induction of metal-binding proteins	Intertidal Tolka estuary
Bile metabolites	Organ	Flatfish: Excretion of chemical metabolites from liver	Subtidal
Comet	Cell	*M. edulis*: Breakdown of genetic DNA	Intertidal Tolka estuary

Ecological Network Analysis (ENA) was adapted from economics for analysis of ecological systems and the impact of changes in trophic transfers in the system have been explored by Ulanowicz (1986, 1997). In this present study, ratios of ENA metrics, being dimensionless, provide a good basic for system assessment (Ulanowicz 1997, Wilson et al. 2007). The two shown here are the diversity of the flows (Connectivity/Throughput, C/T) which provides a measure of the evenness of the trophic links (*pace* Pielou's evenness, J = H'/ H_{max}) and the Finn Cycling Index (FCI) which indicates the maturity of the system through its capacity to retain and recycle energy. Data presented here from Wilson and Parkes (1999) and Wilson et al. (2007).
The Biological Quality Index (BQI) and Pollution Load Index (PLI) (Jeffrey et al. 1985, Wilson 2003) are based on zonal metrics of estuarine systems measuring macrofaunal community stage (as per Pearson and Rosenberg 1978) and sediment contaminant levels respectively. The data presented here is from Wilson (2003) and other unpublished data from 1979 – 2010.
All other indices form part of the project *Biological effects and chemical measurements in Irish marine waters* (PBA-ME-07-001) ((Giltrap et al. 2013), based on ICES recommendations (Thain et al. 2008, Davies and Vethaak 2012 and references therein). The full project report contains data on the Dublin sites

and a range of other samples included in spatial comparisons across 6 Irish estuaries (Giltrap et al. 2013). Unless otherwise indicated, fish data refer to samples taken from the same location as the AMBI data, and mussel (*Mytilus edulis*) results from intertidal mussel beds in the Tolka estuary and Bull Lagoon.
The coefficient of variation (CV) is expressed as the perecentage of the standard deviation over the mean.

Results

The metrics from the ENA and the long-term means of the BQI and PLI are shown along with the coefficient of variation in Table 2.

Table 2. System indices, showing number of observations (n), mean and standard deviation (SD) and coefficient of variation (CV, %).

Index	N	Mean	SD	CV
ENA C/T ratio	5	2.59	0.23	8.7%
ENA FCI	5	0.52	0.17	33.1%
PLI	15	0.83	0.44	53.7%
BQI	19	2.44	1.70	69.5%

There was a considerable range in the CV values at system level, with a low CV for the ENA C/T ratio despite the fact that this was based on a fairly small number (5) of analyses. The variability in the BQI and PLI was somewhat higher, although it must be remembered that these are annual readings, in which any trend over time will inevitably increase the variability.

A similar range of variability was seen in the community-level indices (Table 3).

Table 3. Community-level indices: legend as Table 2.

Index	N	Mean	SD	CV
Ambi BC	25	1.45	0.30	20.5%
1-(AMBI/7)	25	0.79	0.042	5.4%
S	25	35.2	17.9	50.8%
Shannon-Weiner	25	3.85	0.71	18.5%
Evenness, J	25	0.78	0.12	15.9%

Surprisingly perhaps, the highest variability was seen in the number of species (S) rather than in any of the calculated indices, while the formula (1-AMBI/7) gives an artificially-low CV of 5.4%. The other three were remarkably consistent at around 20% (Table 3).

Table 4. Population-level indices: legend as Table 2.

Index	N	Mean	SD	CV
SOS residuals (probits)	20	0.075	0.05	67.5%

Mortality (%)	9	11.4	33.2	290%
Flatfish Liver Abnormalities (%)	3	67.3	6.98	10.4%
Liver NSL	3	29.1	6.59	22.6%
Liver FCA	3	31.3	5.44	17.4%
Imposex VDSI>2	7	27.1	22.5	83.0%

Variability in the mortality indices (Table 4) was very high, reflecting the fact that, while the great majority (7/9) indicated no mortality at all, there was 100% mortality in the *Skeletonema* assay.

Table 5. Individual-level indices: legend as Table 2.

Index	N	Mean	SD	CV
Fulton's CF	12	0.71	0.13	18.9%
SFG ($J.h^{-1}.g^{-1}$)	12	4.38	2.36	53.9%
Respiration ($J.h^{-1}.g^{-1}$)	12	5.50	1.60	29.0%
Clearance rate ($l.h^{-1}.g^{-1}$)	12	1.98	0.30	15.2%

Both respiration, and clearance rate which are component measures of SFG had smaller CVs than SFG itself, which was by some way the most variable of the individual-level indices (Table 5).

Table 6. Individual-level TBT (EDC) indices: legend as Table 2.

Index	N	Mean	SD	CV
Imposex (all)	145	1.20	1.48	123%
Imposex (Poolbeg)	25	1.56	1.78	113%
Imposex (South Wall)	18	1.84	1.64	89.1%
Intersex (all)	82	0.59	0.57	96.6%
Intersex (South Wall)	32	0.66	0.65	99.5%

The CVs for the EDC indices were all very high (Table 6), whether as taken from all sites combined in Dublin Bay, or taken from the sites closest to the Tolka estuary at Poolbeg and the South Wall.

At cellular level there was again a wide range of CV values (Table 7), with an exceptionally high CV (208%) for the vitellogenin assay in male plaice. Two individuals in this assay had levels of vitellogenin almost two orders of magnitude higher than any other individual, but even omitting those individuals from the calculations still left a very high CV of 143%. In contrast, the CV of vitellogenin in dab was low. There were other interspecific differences in CV in other assays but none were as marked as for vitellogenin (Table 7).

Table 7. Cellular-level indices: legend as Table 2.

Index	N	Mean	SD	CV
FRAP (all) (mM Fe(II).mg protein^{-1})	45	1.86	1.46	78.1%

FRAP (Tolka estuary) (mM Fe(II).mg protein^{-1})	15	0.23	0.089	39.8%
EROD (dab) (pM.min^{-1} .mg protein^{-1})	33	17.8	17.1	95.8%
EROD (plaice) (pM.min^{-1} .mg protein^{-1})	20	29.8	13.1	44.1%
NRR (mins)	72	72.5	41.7	57.6%
ALP (µg.mg protein^{-1}) (dab, M)	9	3.53	1.65	46.6%
ALP (µg.mg protein^{-1}) (*M. edulis)*	30	13.8	15.6	112%
Vitellogenin (µg.ml^{-1}) (dab, M)	17	0.22	0.05	23.9%
Vitellogenin (µg.ml^{-1}) (plaice, M)	10	28.6	59.6	208%
AChE (brain) (nM ACTC.min^{-1} .mg protein^{-1})	21	637	171	26.9%
AChE (muscle) (nM ACTC.min^{-1} .mg protein^{-1})	23	152	40.5	26.6%
AChE (*M. edulis*) (nM ACTC.min^{-1} .mg protein^{-1})	30	77.4	27.2	35.1%
MT (all) (µg.mg protein^{-1})	45	2.24	0.81	36.1%
MT (Tolka estuary) (µg.mg protein^{-1})	15	2.02	1.22	60.0%
Bile metabolites (ppm) (dab)	25	0.20	0.08	39.4%
Bile metabolites (ppm) (plaice)	13	0.22	0.03	14.2%
Comet	40	1.69	1.18	69.9%

Taking all the data in Tables 2 to 7, the variability in each of the indices was plotted against the level of organization from 5 (system-level) to 1 (cell level) (Figure 3).

There was no significant relationship (Figure 3) of variability with level of organization either with ($R^2 = 0.01$, $p = 0.27$) or without ($R^2 = 0.065$, $p = 0.06$) the very high CV for the (level 3) mortality tests (Table 4).

Discussion

There was a wide range of variability at all levels of organization, with the same test (e.g. EthoxyResorufin-O-Deethylase, EROD) exhibiting not only different results but also different variability depending on the species tested.

In terms of the status of Dublin Bay, there is no consistent picture. Analysis of OSPAR (Oslo and Paris Commissions) Coordinated Environmental Monitoring Programme (CEMP) contaminants (Giltrap et al. 2013) found 6 sediment contaminants (out of a total of 30) exceeding the Environmental Assessment Criteria (EAC) with 3 (all metals) below the OSPAR Background Assessment Criteria (BAC), and a similar picture for contaminants in *M. edulis* (2/16 above EAC and 3 below BAC). Thus, while contaminant pressures may be equivocal, Dublin Bay could be argued to present an ideal situation to test the performance of the various ecological status indices.

At system level (Table 2), in which BQI or PLI values <1.0 are indicative of impairment (Jeffrey et al. 1985, Wilson 2003), the BQI suggests that the Tolka estuary, while not pristine (BQI = 10) would nevertheless fall into an 'adequate' category, while the PLI ($\bar{x} < 1$) suggests the opposite ('poor'). The latter assessment is the more surprising since the OSPAR CEMP results suggest a much better, if not totally uncontaminated, status. The PLI assessment is supported to some degree by the ENA metrics, which are at the lower end of the status range reported from elsewhere (Wilson et al. 2007).

All the calculated indices at community (Table 3) and system (Table 2) level had much lower variability than the raw species numbers, and across all the levels of organization, were the category with the lowest CVs (Figure 2). However, some caution is required because of the formulae used, with the addition of a constant in the [1-AMBI/7] index masking the true variability in this measure. Contrary to the system-level indices, these assessments suggested good status with the AZTI Marine Biological Index Biotic Coefficient (AMBI BC) suggesting a system of just under the highest (BC < 1.2) quality (Borja et al. 2003).

Likewise the population-level measures (Table 4) suggested good to fair quality. The LT_{50} for the Stress-on-Stress (SOS) was 13.3 days (95% confidence interval 13.1-13.5) which is well above the OSPAR BAC of 10 days. There was zero mortality in almost all of the sediment toxicity tests, except for *S .costatum*, in which all 3 replicates yielded 100% mortality. These tests presented some difficulty in assessing variability, since, when the individual species' tests were considered in isolation, the CV was (with 1 exception) zero, since the 3 replicates were all the same (0% or 100%). The data in Table 4 is the variability across the suite of recommended tests, justified by the different time scales set in the standard protocols. The removal of this value from the variability/level comparison (Figure 2) did strengthen the relationship, although still not quite to a statistically-significant ($p<0.05$) level. The fish pathology results suggested a rather lower status, with for example the incidence (31.3%) of hepatic foci of cellular alteration (FCA) at a level to raise concern (Stentiford et al. 2003).

The contrast at individual level between the high level of variability in all the highly-specific TBT responses (Table 6) and the generalized stress responses (Table 5) was marked, and this can be clearly seen in Figure 2. That there should be such high variation in imposex and intersex is the more unexpected, since the former indicated that in the Bay in general, and even in the sites closest to the port and harbor the mean Vas Deferens Status Index (VDSI) scores were below OSPAR EAC (VDSI > 2.0) although well above the OSPAR BAC (VDSI < 0.3) and perhaps suggesting the residues from a past history of TBT contamination rather than much present contamination. The mean SFG was less than that set for the OSPAR EAC (5 $J.h^{-1}.g^{-1}$), indicating only 'fair' status. One animal actually registered a negative Scope-for-Growth (SFG) (which is clearly unsustainable over time) and this range of values is reflected in the relatively high CV (Table 5).

As the SFG components (Respiration, R, and Clearance Rate, CR) had much lower CVs, the variability may clearly be additive.

A wide variety of cell-level assays were tested and there was a great deal of variation among and even within the different tests (Table 7). Those, such as the dab EROD, AchE and bile metabolites are all within the OSPAR EAC category indicating that the Bay is of good ecological status. The variability however must raise concern not just in terms of a consistent status rating, but also for the reliability and consistency of the individual assays themselves, even for those such as EROD for which reference values have been put forward. The results for vitellogenin provide a good example of the lack of reliability. While the levels for dab were reasonably consistent (Table 7), even if they did suggest, along with the ALP assays, some concern that male dab in Dublin Bay were becoming feminised, those for plaice were wildly variable, and the influence on the variability-level correlation can be clearly seen (Figure 3).

The greater the variability in an indicator, the more difficult it is to make a defintive judgement on status. This applies not only to extreme cases, such as the plaice vitellogenin or the sediment toxicity tests discussed above, but also to any which are close to the category boundaries. For example, the 95% confidence limit for SFG (Table 5) lies well above the OSPAR EAC, moving the status classification up one category, while the complete reverse is true for imposex at South Wall (Table 6). Such a consequence for decision making, especially under the "One-Out, All-Out" (OOAO) principle as the suggested approach to ensure conservative actions has been recognised, and the alternative of 'weight of evidence' or fuzzy inference system based on an integrated suite of indicators has been proposed as a more realistic and practical alternative (Gottardo et al. 2011).

Even using only those assays for which there are set or suggested, criteria (OSPAR 2009, SGIMC 2011), Dublin Bay clearly fails on the OOAO principle (Table 8). However, the weight of evidence approach is more balanced with, if anything slightly more evidence for BAC status than for exceeding EAC.

Table 8. Weight of evidence approach for status of Dublin Bay. See also text for comments.

	Fail (> EAC)	Pass	Pass + (<BAC)	No criterion
Chemical Assays	2	11	3	0
Biological Assays	1*	4	11	15

*excluding PLI

The sole biological assay exceeding EAC was SFG, but, unlike the chemical criteria, a large proportion of the biological assays met the BAC (Table 8). For the specific assays, it was noticeable that some polycyclic Aromatic Hydrocarbond (PAHs, e.g. Phenanthrene) were <BAC as was the Dab EROD assay, while in contrast, some PCB congeners exceeded the EAC, yet the AchE assay was well below BAC. As almost half the assays presented here have yet to have assigned BAC or EAC values and for those for which BAC and EAC have been put forward

there is still evidently some mis-match between the chemical and the biological assays, so more clearly need to be done.

It might be argued that a positive or negative result from an assay at cellular level need not be necessarily consistent with the result fom the others. For example there is no reason why the presence of female chemicals (ALP, Vitellogenin) in male fish should impinge on their general fitness (FRAP) or their response to specific stressors such as metals (MT), PAHs (EROD) or other xenobiotics (AChE). However, there is a clear chain of consequence that could or should lead to impacts at population level and above, and the EDC TBT provides a classic example of the impacts through the levels of organisation (Hawkins et al. 1994). Nevertheless, Hawkins et al. (1994) still considered that definitive evidence of TBT impact at community level remained to be established, and, although others (e.g. Borja et al. 2010) have been less conservative in their assertions, they also acknowledged the complicating effects of the many other variables at this level.

There is no evidence from the data presented here that monitoring at lower levels of organization reduced the variability of the measure despite the extra control gained over other influencing variables. Nor, despite some slight indication from SFG and its components R and CR, was their evidence that variability was in any way additive up levels of organization. The conclusion must be, therefore, that either homeostatic mechanisms operate up the system, or that there is a degree of variability inherent in any measure, independent of influences such as level of organization and the number of uncontrolled variables.

Ellis (1977) in his classic text on sampling suggested that a CV of 20% would be acceptable for benthic species' monitoring: applying that criterion to the indices presented here leaves relatively few as candidates for an integrated monitoring system. Those clearly failing the 20% CV barrier include several currently recommended including SOS, SFG, imposex and all the cell-level measures, with others (e.g. AMBI BC) on the borderline. While the current protocols specify measures to reduce variability (e.g. restrictions on the size of organism to be used or the season of testing), these are clearly insufficient. It is therefore suggested that a more profitable approach in the long-term would be to establish the sources of the variability and to produce models which would not only account for sources of variation in the test, but also allow it to be used in a predictive manner over a greater range of conditions than at present.

Acknowledgements

The authors would like to acknowledge the support of Seachange grant No. PBA-ME-07-001 with the support of the Marine Institute and the EPA. They would like to thank their many colleagues who assisted in this study and also the anonymous reviewers for their helpful comments and suggestions.

References

Blaise, C., Gagne, F., Pellerin, J. & Hansen, P. D. 1999. Determination of vitellogenin-like properties in Mya arenaria hemolymph (Saguenay Fjord, Canada): a potential biomarker for endocrine disruption. Environmental Toxicology, 14, 455-465.

Bocquené, G. & Galgani, F. 1998. Biological effects of contaminants: cholinesterase inhibition by organophosphate and carbamate compounds. ICES Techniques in Marine Environmental Sciences, 22, 1-12.

Bocquené, G., Galgani, F. & Truquet, P. 1990. Characterization and assay conditions for use of AChe activity from several marine species in pollution monitoring. Marine Environmental Research, 30, 75-89.

Borja, A., I. Muxika, & J. Franco 2003. The application of a marine biotic index to different impact sources affecting soft-bottom benthic communities along European coasts. Marine Pollution Bulletin 46, 835–845.

Borja, A., M. Elliott, J. Carstensen & A-S. Heiskanen 2010. Marine management – Towards an integrated implementation of the European Marine Strategy Framework and the Water Framework Directives. Marine Pollution Bulletin 60, 2175 – 2186.

Davies, I. M. & A.D. Vethaak (eds) 2012. Integrated marine environmental monitoring of chemicals and their effects. ICES Cooperative Research Report No. 315. 277 pp.

Ellis, J.M. 1977. Some methods for the statistical analysis of samples of benthic invertebrates. Freshwater Biological Association Scientific Publication No. 25. Freshwater Biological Association, Ambleside, UK. 159 pp.

EPA 2010. Water Quality in Ireland 2007-2009. Environmental Protection Agency, Johnstown, Co. Wexford, Ireland.

Gibbs, P.E., G.W. Bryan, & P.L. Pascoe 1987. The use of the dog-whelk, Nucella lapillus, as an indicator of tributyltin (TBT) contamination. Journal of the Marine Biological Association of the United Kingdom 67, 507-523.

Giltrap, M., B. McHugh, J. Ronan, J.G. Wilson & E. McGovern. 2013. Biological effects and chemical measurements in Irish marine waters (PBA-ME-07-001). Marine Institute, Oranmore, Co. Galway, Ireland.

Gottardo, S., E. Semenzin, S. Giove, A. Zabeo, A. Critto, D. de Zwart, A. Ginebreda, & A. Marcominia, 2011. Integrated risk assessment for WFD ecological status classification applied to Llobregat river basin (Spain). Part I—Fuzzy approach to aggregate biological indicators. Science of The Total Environment 409: 4701–4712

Gray, J.S. & F.B. Mirza 1979. A possible method for the detection of pollution-induced disturbance on marine benthic communities. Marine Pollution Bulletin 10: 142-146.

Hawkins, S.J., S.V. Proud, S.K. Spence & A.J. Southward, 1994. From the Individual to the community and beyond: water quality, stress indicators and key species in coastal ecosystems. In Sutcliffe, D.W. (Ed) (1994) Water Quality and Stress Indicators in Marine and Freshwater Systems: Linking Levels of Organisation. Freshwater Biological Association, Ambleside, 35 - 62.

ICES 2011. Further development of guidance on integrated monitoring and assessment of chemicals and biological effects. OSPAR special request advice June 2011. ICES Advice, 2011, Book 1: 1.5.5.4.
Irvine, K. 2004. Classifying ecological status under the European Water Framework Directive: the need for monitoring to account for natural variability. Aquatic Conservation: Marine and Freshwater Ecosystems
Jeffrey, D.W., J.G. Wilson, C.R. Harris, & D.L. Tomlinson, 1985. The application of two simple indices to Irish estuary pollution status. In J.G. Wilson and W. Halcrow (eds) *Estuarine management and quality assessment,* 147 – 65. London. Plenum Press.
Leppakoski, E. 1977. Monitoring the benthic environment of organically polluted river mouths. *Biological Monitoring of Inland Fisheries*, 125-132.
OSPAR, 2009. Background Document on CEMP Assessment Criteria for the QSR 2010. Monitoring and Assessment Series. OSPAR Commission, London.
Pearson, T.H. & R. Rosenberg, 1978. Macro- benthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanography and Marine Biology Annual Review* **4,** 481–520.
Porte, C., M. Sole, J. Albaigés, & D.R. Livingstone 1991. Responses of mixed-function oxygenase and antioxidase enzyme system of Mytilus sp. to organic pollution. Comparative Biochemistry and Physiology Part C: Comparative Pharmacology, 100: 183–186.
SGIMC 2011. Report of the Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC), 14–18 March 2011, Copenhagen, Denmark. ICES CM 2011/ACOM:30, 1-265.
Singh, N. P., McCoy, M. T., Tice, R. R. & Schneider, E. L. (1988) A simple technique for quantitation of low levels of DNA damage in individual cells. Experimental Cell Research, 175, 184-191.
Stebbing, A.R.D. 1981. Hormesis‒stimulation of colony growth in Campanularia flexuosa (Hydrozoa) by copper, cadmium and other toxicants. Aquatic Toxicology 1: 227-238.
Stentiford, G.D., M. Longshaw, B.P. Lyons, G. Jones, M. Green & S.W. Feist, 2003. Histopathological biomarkers in estuarine fish species for the assessment of biological effects of contaminants. Marine Environmental Research 55: 137–159.
Thain, J. E., Vethaak, A. D., and Hylland, K. 2008. Contaminants in marine ecosystems: developing an integrated indicator framework using biological-effect techniques. – ICES Journal of Marine Science, 65: 1508–1514.
Thorson, G. 1957. Bottom Communities. In J.W. Hedgepeth (ed.) Treatise on Marine Ecology and Palaeoecology. Geological Society of America Memoir 67: 461-534.
Viarengo, A., Ponzano, E., Dondero, F. & Fabbri, R. (1997) A simple spectophotometric method for metallothionein evaluation in marine organisms: an application to Mediterranean and Antarctic molluscs. Marine Environmental Research, 44, 69-84.
Wilson JG. 2003. Evaluation of estuarine quality status at system level with the Biological Quality Index (BQI) and the Pollution Load Index (PLI). Biology and Environment: Proceedings of the Royal Irish Academy 103B: 47–59.

Wilson, J.G., Rybarczyk, H., Elkaim, B., 2007. A comparison of energy flow through the Dublin Bay and Baie de Somme intertidal ecosystems and their network analysis. Hydrobiologia 588, 231–243.

Glossary of Abbreviations

Abbreviation	
AChE	AcetylCholine Esterase
ALP	Alkaline-Labile Phosphate
AMBI	AZTI Marine Biotic Index
AZTI	Technological Institute for Fisheries and Food, San Sebastian, Spain
BAC	Background Assessment Criteria
BC	Biotic Coefficient
BOD	Biochemical Oxygen Demand
BQI	Biological Quality Index
C/T	Connectivity/Throughput ratio
CEMP	Coordinated Environmental Monitoring Programme
COX	Cytochrome Oxidase
CR	Clearance Rate
CV	Coefficient of Variation
DNA	Deoxyribonucleic acid
EAC	Environmental Assessment Criteria
EDC	Endocrine-Disrupting Compound
ENA	Ecological Network Analysis
EQS	Ecological Quality Standard
EROD	Ethoxyresorufin-O-deethylase
FCA	Foci of Cellular Alteration
FCI	Finn Cycling Index
FRAP	Ferric Reducing Ability of Plasma
H', H_{max}	Shannon-Weiner Index, maximum value
ICES	International Council for the Exploration of the Sea
J	Pielou's Index of Evenness
LMS	Lysosome Membrane Stability
LT_{50}	Lethal Time for 50% effect
MSFD	Marine Strategy Framework Directive
MT	Metallothionein
NRR	Neutral Red Retention
NSL	Non-Specific Lesion
OOAO	One Out All Out
OSPAR	Oslo and Paris Conventions
PAH	Polycyclic Aromatic Hydrocarbon
PCB	PolyChlorinated Biphenyl
PLI	Pollution Load Index
R	Respiration
S	Number of Species
SFG	Scope for Growth
SGMIC	Study Group on Integrated Monitoring of Contaminants and Biological Effects
SOS	Stress-On-Stress
TBT	Tri-Butyl Tin
VDSI	Vas Deferens Sequence Index

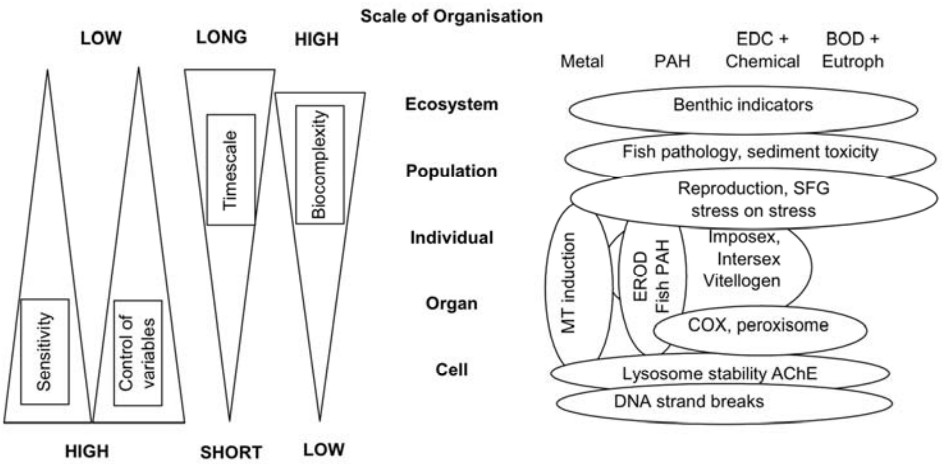
Vtg	Vitellogenin
WFD	Water Framework Directive

Legends for Figures.

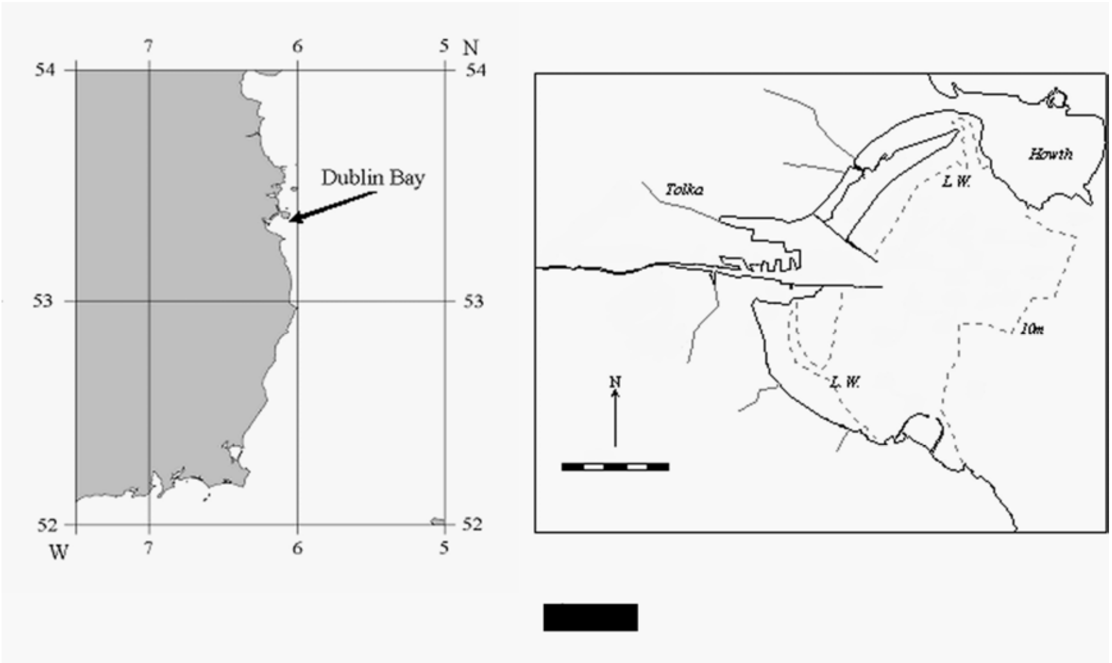
Figure 1. Summary of index properties against scale of organization. See also text for explanation and discussion.

Figure 2. Dublin Bay showing Liffey and Tolka estuaries and extent of the littoral area (dotted line).

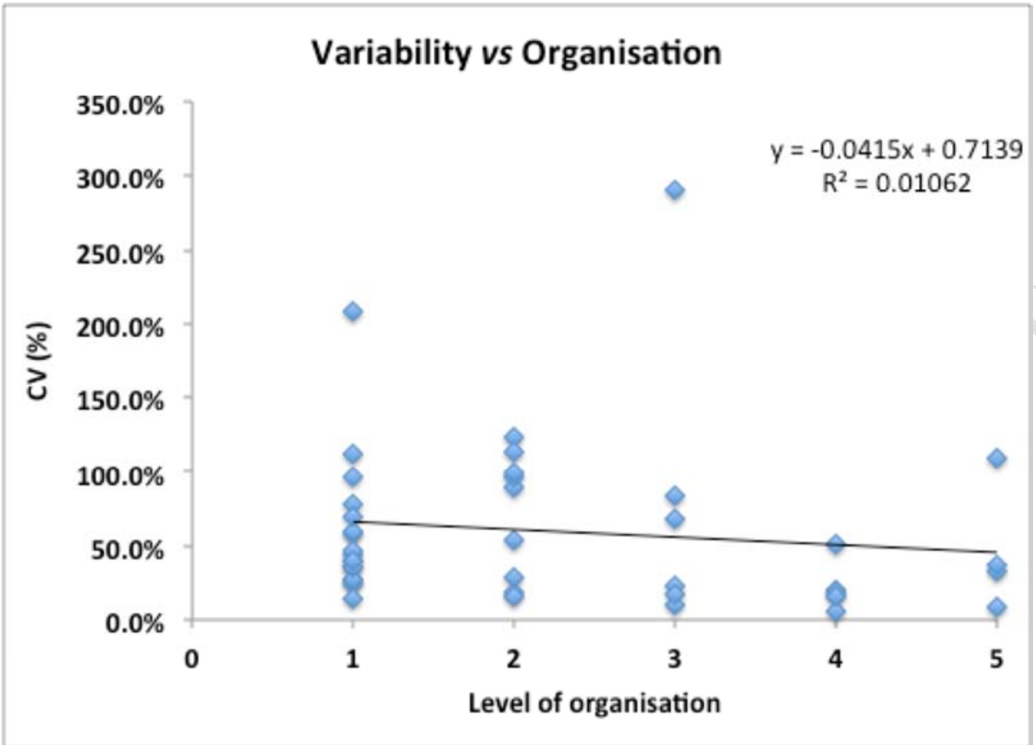
Figure 3. Variability (CV%) against level of organization from 1 (Cell) to 5 (System); n = 40.



478
479
480
481
482 Wilson et al. Fig. 1.
483
484



Wilson et al. Fig. 2



Wilson et al. Fig 3.

Highlights (Wilson et al)

- Integrated suite of biomarker assessments
- Comparison of diagnostic consistency across biomarkers
- Evaluation of variability and control with ecological organisation